Periods of Iterated Rational Functions over a Finite Field

Charles Burnette
Department of Mathematics
Drexel University
Philadelphia, PA 19104-2875
cdb72@drexel.edu

Eric Schmutz
Department of Mathematics
Drexel University
Philadelphia, PA 19104-2875
Eric.Jonathan.Schmutz@drexel.edu

June 28, 2016

Abstract

If f is a polynomial of degree d in $\mathbb{F}_q[x]$, let $c_k(f)$ be the number of cycles of length k in the directed graph on $\mathbb{F}_q$ with edges $\{(v, f(v))\}_{v\in\mathbb{F}_q}$. For random polynomials, the numbers $c_k, 1 \le k \le b$, have asymptotic behavior resembling that for the cycle lengths of random functions $f : [q] \to [q]$. However random polynomials differ from random functions in important ways. For example, given the set of cyclic (periodic) points, it is not necessarily true that all permutations of those cyclic points are equally likely to occur as the restriction of f. This, and the limitations of Lagrange interpolation, together complicate research on $\mathbf{T}(f) =$ the ultimate period of f under compositional iteration. We prove a lower bound for the average value of $\log \mathbf{T}$: if $d = d(q) \to \infty$, but $d = o(\sqrt{q})$, then the expected value of $\log \mathbf{T}$ is

$$\mathbb{E}(\log \mathbf{T}) := \frac{1}{q^d(q-1)} \sum_f \log \mathbf{T}(f) > \frac{d}{2}(1 + o(1)),$$

where the sum is over all $q^d(q-1)$ polynomials of degree d in $\mathbb{F}_q[x]$. Similar results are proved for rational functions.

1 Introduction

The classical theory of "random mappings"concerns functions chosen randomly from among the q^q functions whose domain and codomain are a given q-element set V. (See, for example, [8, 10, 14].) Basic facts about random mappings have motivated conjectures and theorems in "arithmetic dynamics" [2, 3, 16, 18, 20, 21]. This paper includes $\mathbb{F}_q[x]$-analogs of known facts about random mappings' cycle lengths and the lengths of their (ultimate) periods under function composition.

If f is any function from $\mathbb{F}_q$ to $\mathbb{F}_q$, let $\textsc{cyclic}(f)$ be the set of periodic points, i.e the set of vertices that lie on cycles in the "functional digraph" G_f that is formed by putting an edge from v to $f(v)$ for each v. Let $\mathbf{Z}(f) = |\textsc{cyclic}(f)|$ and let σ_f be the restriction of f to $\textsc{cyclic}(f)$. Finally, let $\mathbf{T}(f)$ be the least common multiple of the cycle lengths, i.e. the order of σ_f as an element of the group $\mathrm{Sym}(\textsc{cyclic}(f))$. Equivalently, $\mathbf{T}(f)$ is the ultimate period of f, defined as the smallest positive integer t such that, for every $n \geq q$, $f^{(n+t)} = f^{(n)}$.

A theorem of Harris [11] tells us how large $\mathbf{T}$ normally is for random mappings: if $\epsilon > 0$, then for all but $o(q^q)$ functions f,

$$e^{(\frac{1}{8}-\epsilon)\log^2 q} < \mathbf{T}(f) < e^{(\frac{1}{8}+\epsilon)\log^2 q}. \tag{1.1}$$

The average value of $\mathbf{T}$ is much larger [19]. Igor Shparlinski and Alina Ostafe proposed the challenging problem of estimating the average value of $\mathbf{T}$ in the much harder case of degree d polynomials and rational functions. We require that d be less than q, because otherwise all q^q mappings can be realized by Lagrange interpolation as degree d polynomials, and we are back to random mappings.

Let $\Omega(q,d)$ be the set of all $q^d(q-1)$ polynomials of degree d with coefficients in the finite field $\mathbb{F}_q$. Let $\mathbb{M} = \mathbb{M}_{q,d}$ be the uniform probability measure on $\Omega(q,d)$; for all $A \subseteq \Omega(q,d)$, $\mathbb{M}(A) = \frac{|A|}{q^d(q-1)}$. Then $\mathbf{T}$, $\log \mathbf{T}$, $\mathbf{Z}$, and any other quantities of interest may be regarded as a random variables on $\Omega(q,d)$, and the theorems of combinatorial probability are applicable. In particular, by grouping together elements with the same period, we see that the average value of $\mathbf{T}$ is just the expected value $\mathbb{E}_{\mathbb{M}}(\mathbf{T})$:

$$\frac{1}{q^d(q-1)} \sum_{f \in \Omega(q,d)} \mathbf{T}(f) = \sum_{j \geq 0} j \mathbb{M}(\{f : \mathbf{T}(f) = j\}) = \mathbb{E}_{\mathbb{M}}(\mathbf{T}).$$

Here we prove, using elementary methods, that

$$\frac{1}{q^d(q-1)} \sum_{f \in \Omega(q,d)} \log(\mathbf{T}(f)) \geq \sum_{p \leq \frac{d}{2}}^{*} \frac{\log p}{p} \prod_{j=0}^{p-1}(1 - \frac{j}{q}) - \sum_{p \leq \frac{d}{2}}^{*} \frac{\log p}{2p^2} \prod_{j=0}^{2p-1}(1 - \frac{j}{q}),$$

where the $*$ indicates that the sum is restricted to prime numbers. As one consequence, we deduce the following crude lower bound for the average value of $\mathbf{T}$ as an incomplete answer to Shparlinski and Ostafe's question. If $d = d(q) \to \infty$ but $d = o(\sqrt{q})$, then

$$\frac{1}{q^d(q-1)} \sum_{f \in \Omega(q,d)} \mathbf{T}(f) \geq (1 + o(1))\frac{d}{2}.$$

In some ways, it is easier to work with random mappings than with random polynomials. In both Harris' proof [11], and in subsequent work on the average period [19], essential use was made of the following observation about classical random mappings and conditional probabilities: given the set of cyclic vertices, all permutations of those vertices are equally likely to occur as the restriction of f. A more formal statement in terms of conditional probabilities is

Observation 1.1. (Folklore) Suppose $\mathbb{M}^*$ is the uniform probability measure on the $|V|^{|V|}$ functions $f : V \to V$. Then, for all non-empty $A \subseteq V$ and all permutations σ of A,

$$\mathbb{M}^*\left(\sigma_f = \sigma \,|\, \textsc{cyclic}(f) = A\right) := \frac{\mathbb{M}^*\left(\sigma_f = \sigma\right)}{\mathbb{M}^*(\textsc{cyclic}(f) = A)} = \frac{1}{|A|!}.$$

However, a simple counting argument shows that the analogue of Observation 1.1 cannot be true in general for $\Omega(d, q)$. Consider, for example, the case where A is all of $\mathbb{F}_q$. If it were true that $\mathbb{M}\left(\sigma_f = \sigma \,|\, \textsc{cyclic}(f) = A\right) = \frac{1}{|A|!} > 0$ for all σ in $\mathrm{Sym}(A)$, then, for each of the $|A|!$ permutations σ, $\Omega(q, d)$ would contain at least one element f for which $\sigma_f = \sigma$. This implies that $|\Omega(q, d)| \geq |A|!$. But if A is all of $\mathbb{F}_q$, and if $d = o(q)$, then $|\Omega(q, d)| < |A|!$ for all sufficiently large prime powers q.

Nevertheless, σ_f is always a permutation of $\textsc{cyclic}(f)$. By an old theorem of Landau, [15], [17], the maximum order that a permutation of an m-element set can have is $e^{\sqrt{m \log m}(1+\epsilon(m))}$, where $\epsilon(m) \to 0$. However we do not know enough about the distribution of $\mathbf{Z}$ to draw any conclusions about $\mathbb{E}_{\mathbb{M}}(\mathbf{T})$. To date, we have only the trivial upper bound $\mathbb{E}_{\mathbb{M}}(\mathbf{T}) \leq e^{\sqrt{q \log q}(1+o(1))}$ that follows from $\mathbf{Z} \leq q$.

2 Small Cycle Lengths.

Flynn and Garton [9] used Lagrange interpolation to calculate the expected number of cycles of length k. In this section, Flynn and Garton's methods are combined with the method of factorial moments to determine the asymptotic behaviour of the cycle length multiplicities $c_1, c_2, \ldots, c_b$ for fixed b. As we shall see, the cycle lengths are not independent, but they are asymptotically independent as $q \to \infty$. They behave very much like the cycle lengths of random permutations and random mappings.

Lemma 2.1. *Let $A \subset \mathbb{F}_q$, and let σ be a permutation of A. If $d \geq |A|$, then there are exactly $q^{d-|A|}(q-1)$ polynomials in $\mathbb{F}_q[x]$ of degree d that extend σ to $\mathbb{F}_q$.*

Proof. Let $\mathcal{E}_d$ consist of those polynomials f of degree d that extend σ, i.e. such that $f(x) = \sigma(x)$ for all $x \in A$. Also let $\mathcal{K}_d$ be the set of polynomials of degree d such that $f(x) = 0$ for all $x \in A$. By the Lagrange interpolation theorem, $f \in \mathcal{E}_d$ if and only if

$$f(x) = L(x) + k(x),$$

where L is the minimum degree interpolating polynomial (whose degree is less than d), and $k(x) \in \mathcal{K}_d$. Since $\mathbb{F}_q[x]$ is a Euclidean domain, a polynomial of degree d is in $\mathcal{K}_d$ if and only if it is divisible by $g(x) = \prod_{a \in A}(x - a)$. Thus $k \in \mathcal{K}_d$ if and only if there is a polynomial h of degree $d - |A|$ such that $k(x) = g(x)h(x)$. The number of ways to choose the polynomial h is $q^{d-|A|}(q-1)$. Therefore $|\mathcal{E}_d| = |\mathcal{K}_d| = q^{d-|A|}(q-1)$. □

Before proving the next theorem, we review some basic facts from probability. Many readers will be familiar with the Method of Moments. A glib summary is that, to prove that

a sequence of random variables X_n converges to X, it suffices (under mild conditions) to prove that, for each r, the expected value of X_n^r converges to the expected value of X^r. For multivariable extensions, it is necessary to consider cross moments: to prove that (X_n, Y_n) converges to (X, Y), one proves that $\mathbb{E}(X_n^{r_1} X_n^{r_2})$ converges to $\mathbb{E}(X^{r_1} Y^{r_2})$ for non-negative integers r_1, r_2. Section 30 of Billingsley [5] contains a more precise and detailed presentation of the Method of Moments, together with a number-theoretic application. Note that each of the powers X^r can be written as a linear combination of falling the falling factorials

$$(X)_s := \prod_{j=0}^{s-1} (X - j), s = 1, 2, \ldots, r$$

and vice versa. Hence there is a corresponding "Method of Factorial Moments"; convergence of the factorial moments implies convergence of the moments, and vice versa. This method is especially convenient for arithmetic and enumerative applications because $(X)_r$ occurs naturally as the number of injective functions from an r-element set to an X-element set. It is what George Andrews ([1], page 32) refers to as the number of r-permutations of an X element set. If X_1, X_2 have a Poisson distributions with parameters λ_1, λ_2, i.e. if $\Pr(X_i = m) = e^{-\lambda_i} \frac{\lambda_i^m}{m!}, i = 1, 2; m = 0, 1, 2, \ldots$, then for any r the r'th factorial moment of X_i is λ_i^r, i.e. $\mathbb{E}((X)_r) = \lambda_i^r$. With the additional hypothesis that X_1 and X_2 are independent, the falling factorials are also independent and the expectation of their product is the product of their expectations: $\mathbb{E}((X)_{r_1}(X)_{r_2}) = \lambda_1^{r_1} \lambda_2^{r_2}$. Section 6.1 of [13] is directly relevant to our application of the Method of Factorial Moments. It includes (bottom of page 144) a generalization of Lemma 2.2 below.

Before stating the lemma, we fix some notation. If σ is a permutation of a subset of $\mathbb{F}_q$, define an "indicator"(i.e. characteristic function) by $I_\sigma(f) = 1$ if σ is the restriction of f to the set of elements that σ permutes, and $I_\sigma(f) = 0$ otherwise. Let $\mathcal{Z}_k$ be the set of all k-cycles that can be formed using elements of $\mathbb{F}_q$. Then

$$c_k(f) = \sum_{C \in \mathcal{Z}_k} I_C(f)$$

is the number of cycles of length k that f has. Also note that a product of indicators is itself an indicator with a concrete interpretation: $I_{C_1} I_{C_2} \cdots I_{C_r}(f)$ is one if and only if the r cycles $C_1, C_2, \ldots, C_r$ are all cycles of f, i.e $I_{C_1} I_{C_2} \cdots I_{C_r}(f) = I_\sigma(f)$, where σ is the permutation having cycles $C_1, C_2, \ldots, C_r$.

Lemma 2.2. *With the notation $I_C(f) = 1$ if C is a k-cycle for f, and zero otherwise, we have*

$$c_k(c_k - 1) \cdots (c_k - r + 1) = \sum I_{C_1} I_{C_2} \ldots I_{C_r},$$

where the sum is over all sequences of r disjoint k-cycles. In other words, there are exactly $(c_k(f))_r$ ways to pick a sequence of r disjoint k cycles of f.

Recall that $\mathbb{M}(A) = \frac{|A|}{q^d(q-1)}$ for all sets A of degree d polynomials.

Theorem 2.3. *If $d = d(q) \to \infty$, then for any fixed b the random variables c_k, $k = 1, 2, \ldots, b$, are asymptotically independent Poisson($1/k$). In other words, for for any non negative integers $m_1, m_2, \ldots, m_b$,*

$$\lim_{q\to\infty} \mathbb{M}(c_k = m_k, 1 \le k \le b) = \prod_{k=1}^{b} e^{-\frac{1}{k}} \frac{1}{m_k! k^{m_k}}.$$

Comment: The conclusion might well be true when d is fixed. However the hypothesis $d(q) \to \infty$ is needed for our proof because of the way Lagrange interpolation is used.

Proof. By [Theorem 6.10,13] (or [Theorem 21,6]), it suffices to show that the joint factorial moments $\mathbb{E}_{\mathbb{M}}\left((c_1)_{r_1}(c_2)_{r_2}\cdots(c_b)_{r_b}\right)$ converge to those of the corresponding independent Poisson distributions. In other words, it suffices to check that, for any choice of $r_1, r_2, \ldots, r_b$,

$$\begin{aligned}\lim_{q\to\infty} \mathbb{E}_{\mathbb{M}}\left(\prod_{k=1}^{b}(c_k)_{r_k}\right) &= \prod_{k=1}^{b}(r_k\text{'th factorial moment of Poisson(1/k) random var.}) \\ &= \prod_{k=1}^{b}\frac{1}{k^{r_k}}.\end{aligned}$$

Let $\Lambda_{\vec{r}}$ denote a given choice of an integer partition with r_k parts of size k, and let $m = \sum_k kr_k$ be the number that $\Lambda_{\vec{r}}$ partitions. In the product $\prod_{k=1}^{b}(c_k)_{r_k}$, we can apply Lemma 2.2 to each factor $(c_k)_{r_k}$:

$$\prod_{k=1}^{b}(c_k)_{r_k} = \prod_{k=1}^{b}(\sum I_{C_{k,1}} I_{C_{k,2}} \cdots I_{C_{k,r_k}}). \tag{2.1}$$

If we now expand the right hand side, there are an unpleasantly large number of terms. On the other hand, each term is nothing more than a product of indicators. We have $I_C I_{C'} = 0$ whenever two cycles C,C' are not disjoint. Therefore, when (2.1) is expanded, each non-zero term in the sum corresponds to a permutation σ of of m field elements having type $\Lambda_{\vec{r}}$ (i.e. having r_k cycles of length k for $1 \le k \le b$). Because there are $r_k!$ possible ways to order the cycles of length k, the indicator $I_\sigma = \prod_{k=1}^{b}\prod_{j=1}^{r_k} I_{C_{k,j}}$ occurs $\prod_k r_k!$ times. Combining the terms with same permutation we get

$$\prod_{k=1}^{b}(\sum I_{C_{k,1}} I_{C_{k,2}} \cdots I_{C_{k,r_k}}) = (\prod_k r_k!)\sum_\sigma I_\sigma, \tag{2.2}$$

where the sum is over all permutations of type $\Lambda_{\vec{r}}$ of m elements of $\mathbb{F}_q$. Averaging over all polynomials in $\Omega(q, d)$, and using the fact that expectation $\mathbb{E}(-)$ is linear (with no assumption of independence), we get

$$\mathbb{E}_{\mathbb{M}}(\prod_{k=1}^{b}(c_k)_{r_k}) = (\prod_k r_k!)\sum_\sigma \mathbb{E}_{\mathbb{M}}(I_\sigma). \tag{2.3}$$

Because $d(q) \to \infty$, whereas σ is fixed permutation, we can invoke Lemma 2.1 to get

$$\mathbb{E}_{\mathbb{M}}(I_\sigma) = \mathbb{M}(\{f \text{ extends } \sigma\}) = \frac{q^{d-m}(q-1)}{q^d(q-1)} = q^{-m}$$

for all sufficiently large q. Thus, in (2.3), all the terms in the sum are equal. There are $\binom{q}{m}$ ways to choose the m elements that σ permutes. By a well known theorem of Cauchy (see, for example, page 18 proposition 1.3.2 of Stanley [22]), an m-element set has exactly $\frac{m!}{\prod_\ell \ell^{r_\ell} r_\ell!}$ permutations with the given partition $\Lambda_{\vec{r}}$ Therefore, for all sufficiently large q,

$$\mathbb{E}_{\mathbb{M}}(\prod_{k=1}^{b}(c_k)_{r_k}) = \prod_k r_k! \cdot \binom{q}{m} \frac{m!}{\prod_\ell \ell^{r_\ell} r_\ell!} q^{-m} = \frac{(q)_m}{q^m} \prod_k \frac{1}{k^{r_k}}.$$

Because m is fixed,

$$\lim_{q\to\infty} \frac{(q)_m}{q^m} = \lim_{q\to\infty} \prod_{j=0}^{m-1}(1 - \frac{j}{q}) = 1.$$

Therefore

$$\lim_{q\to\infty} \mathbb{E}_{\mathbb{M}}(\prod_{k=1}^{b}(c_k)_{r_k}) = \prod_{k=1}^{b} \frac{1}{k^{r_k}}$$

as was to be shown. □

3 Bounds for $\mathbb{E}_{\mathbb{M}}(\log \mathbf{T})$, $\mathbb{E}_{\mathbb{M}}(\mathbf{T})$, and the Typical Period.

Recall that, for $f \in \Omega(q,d)$, $\mathbf{T}(f)$ is the least common multiple of the cycle lengths, and c_k is the number of cycles of length k. The goal in this section is to prove lower bounds for the expected values of $\mathbf{T}$ and $\log \mathbf{T}$ as well as a lower bound for $\mathbf{T}$ that holds with asymptotic probability one.

For any any choice of ξ, the period $\mathbf{T}$ is at least as large as the product (without multiplicity) of prime cycle lengths in the interval $[2, \xi]$. Define $\delta_k(f) = 1$ if $c_k > 0$, and $\delta_k(f) = 0$ otherwise. Then

$$\mathbf{T}(f) \geq \prod_{p\leq\xi}^{*} p^{\delta_p}, \tag{3.1}$$

where the $*$ indicates that the product is restricted to primes. Since $e^x > x$ for all $x \geq 0$, it is clear that $\mathbf{T}(f) = e^{\log \mathbf{T}(f)} \geq \log \mathbf{T}(f)$. Combining this with (3.1), and averaging over degree d polynomials, we get

$$\mathbb{E}_{\mathbb{M}}(\mathbf{T}) \geq \mathbb{E}_{\mathbb{M}}(\log \mathbf{T}) \geq \sum_{p\leq\xi}^{*} \mathbb{E}_{\mathbb{M}}(\delta_p) \log p. \tag{3.2}$$

We postpone the choice of ξ, but note that the bound can only improve if ξ is increased. However our ability estimate $\mathbb{E}_{\mathbb{M}}(\delta_p)$ will depend on d being larger than 2ξ, so the price for a better estimate is that the polynomials must have larger degree.

In order to calculate $\mathbb{E}_{\mathbb{M}}(\delta_p)$, we need to calculate the cardinality of $\mathcal{Z}_p$, the set of all possible p-cycles. There are $\binom{q}{p}$ ways to choose p elements of $\mathbb{F}_q$ to form a p-cycle. As a special case of the aforementioned theorem of Cauchy, there $(p-1)!$ ways to form a p-cycle from p elements (There are $p!$ ways to write down the p elements in a cycle. This overcounts by a factor p because, calculating subscripts $\mod p$, we have $(x_0, x_1, \ldots, x_{p-1}) = (x_i, x_{i+1}, \ldots, x_{i+p})$ for $0 \leq i < p$). Therefore

$$|\mathcal{Z}_p| = (p-1)!\binom{q}{p}. \tag{3.3}$$

To estimate the quantity $\mathbb{E}_{\mathbb{M}}(\delta_p)$ in equation (3.2), note that $|\Omega(q,d)|\mathbb{E}_{\mathbb{M}}(\delta_p)$ is the number of polynomials in $\Omega(q,d)$ that have at least one cycle of length p. This number can be estimated using inclusion-exclusion and Bonferroni inequalities (See, for example, equation (7) page 66 of [22]. For each possible p-cycle C, having C as a cycle is a property that a polynomial may have.) One option is to use cardinalities of sets in the formulae, and divide by $|\Omega(q,d)|$ at the end to get $\mathbb{E}_{\mathbb{M}}(\delta_p)$. Another equivalent option is to work directly with the probabilities as weights.) If C is a cycle, let A_C be the event that f has C as a cycle; $A_C = \{f \in \Omega(q,d) : I_C(f) = 1\}$. Also define

$$S_r = S_r(p) = \sum_{\{C_1, C_2, \ldots, C_r\}} \mathbb{M}\left(\bigcap_{i=1}^{r} A_{C_i}\right) = \sum_{\{C_1, C_2, \ldots, C_r\}} \mathbb{E}_{\mathbb{M}}(\prod_{i=1}^{r} I_{C_i}),$$

where the sums are over all *unordered* r element subsets of $\mathcal{Z}_p$. If p is prime, then by inclusion-exclusion and the alternating inequalities(see, for example, page 91 of [22]), we have $\mathbb{E}_{\mathbb{M}}(\delta_p) = \sum_{j \geq 1} (-1)^{j+1} S_j$, and for any m,

$$\sum_{j=1}^{2m} (-1)^{j+1} S_j \leq \mathbb{E}_{\mathbb{M}}(\delta_p) \leq \sum_{j=1}^{2m-1} (-1)^{j+1} S_j.$$

In particular, with $m = 1$ we get a convenient lower bound:

$$S_1 - S_2 \leq \mathbb{E}_{\mathbb{M}}(\delta_p) \leq S_1. \tag{3.4}$$

So long as $d \geq p$, we can apply Lemma 2.1 to each cycle C in $\mathcal{Z}_p$:

$$\mathbb{M}(A_C) = \frac{q^{d-p}(q-1)}{q^d(q-1)} = q^{-p}.$$

Combining this with our formula (3.3) for the cardinality of $\mathcal{Z}_p$, we get an exact formula for S_1 :

$$S_1 = \sum_{C \in \mathcal{Z}_p} q^{-p} = \binom{q}{p} \frac{(p-1)!}{q^p} = \frac{1}{p} \frac{(q)_p}{q^p} = \frac{1}{p} \prod_{j=0}^{p-1} (1 - \frac{j}{q}). \tag{3.5}$$

Similarly, if $d \geq 2p$, we can apply Lemma 2.1 to any permutation that consists of two disjoint p-cycles. (As before, intersecting cycles contribute zero). Therefore $S_2 = \sum_{\{C_1,C_2\}} q^{-2p}$, where the sum is restricted to disjoint pairs of cycles. For the number of ways to choose a set of two disjoint p-cycles, we again have $\binom{q}{2p}$ choices of field elements to permute. Again, by Cauchy's theorem, there are $\frac{(2p)!}{2p^2}$ permutations of of those elements that have two $p-$cycles. Thus

$$S_2 = \frac{(2p)!}{2p^2}\binom{q}{2p}q^{-2p} = \frac{1}{2p^2}\prod_{j=0}^{2p-1}(1-\frac{j}{q}). \tag{3.6}$$

Putting (3.5) and (3.6) into (3.4) and then (3.2), we get

Theorem 3.1. *If $d \geq 2\xi$, then*

$$\mathbb{E}_{\mathbb{M}}(\log \mathbf{T}) \geq \sum_{p\leq\xi}^{*}\frac{\log p}{p}\prod_{j=0}^{p-1}(1-\frac{j}{q}) - \sum_{p\leq\xi}^{*}\frac{\log p}{2p^2}\prod_{j=0}^{2p-1}(1-\frac{j}{q}),$$

where the $*$ indicates that the sum is restricted to prime numbers.

Clearly the second sum in Theorem 3.1 is bounded by an absolute constant, regardless of how large d and ξ are, since the unrestricted series $\sum_{n=1}^{\infty}\frac{\log n}{n^2}$ is convergent. For the main term, one can make various asymptotic approximations, depending on the choice of d and ξ. For example, if $d = d(q) \to \infty$ but $d = o(\sqrt{q})$, then we can choose $\xi = \frac{d}{2}$ and use the approximation $\log(1-u) = O(u)$ to simplify the product:

$$\exp(\sum_{j=1}^{p-1}\log(1-\frac{j}{q})) = \exp(\sum_{j=1}^{p-1}O(\frac{j}{q})) = e^{o(1)} = 1 + o(1). \tag{3.7}$$

It is well known fact ([12], page 89), reportedly due to Mertens, that $\sum_{p\leq\xi}^{*}\frac{\log p}{p} = \xi + O(1)$ as $\xi \to \infty$. Therefore we have

Corollary 3.2. *If $d = d(q) \to \infty$, but $d = o(\sqrt{q})$,then*

$$\mathbb{E}_{\mathbb{M}}(\log \mathbf{T}) \geq (1+o(1))\frac{d}{2}.$$

Corollary 3.2 does not necessarily mean that most polynomials have period greater $e^{(\frac{1}{2}-\epsilon)d}$. Without more information, one cannot even rule out the existence of a constant upper bound κ such that $\mathbf{T} \leq \kappa$ for all but $o(q^d(q-1))$ polynomials in $\Omega(q,d)$. However the next proposition shows that, in fact, most polynomials have order larger than $\sqrt{\frac{d}{2}}$.

Proposition 3.3. *If $d = d(q) \to \infty$ but $d = o(\sqrt{q})$, then all but $o(q^d(q-1))$ polynomials in $\Omega(q,d)$ have at least one cycle whose length is in the interval $J = [\beta, \beta^2]$, where $\beta(q) = \sqrt{\frac{d}{2}}$.*

Proof. For $f \in \Omega(q,d)$, let $\mathbf{N}(f) = \sum_{k \in J} c_k$ be the number of cycles of f that have length in J. The goal is to show $\mathbb{M}(\mathbf{N} = 0) = o(1)$. If $\mathbf{N} = 0$, then obviously $|\mathbf{N} - \mu| \geq \mu$ for any real number μ. Therefore

$$\mathbb{M}(\mathbf{N} = 0) \leq \mathbb{M}(|\mathbf{N} - \mu| \geq \mu). \tag{3.8}$$

A standard approach is to set $\mu := \mathbb{E}_{\mathbb{M}}(\mathbf{N})$, and $\sigma^2 := \mathbb{E}_{\mathbb{M}}((\mathbf{N} - \mu)^2)$, and use Chebyshev's inequality(see, for example, page 75 of [5]) to show that the right side of (3.8) approaches zero. By Chebyshev's inequality,

$$\mathbb{M}(|\mathbf{N} - \mu| \geq \mu) \leq \frac{\sigma^2}{\mu^2}. \tag{3.9}$$

It therefore suffices to show that $\sigma^2 = o(\mu^2)$. By elementary algebra,

$$(\mathbf{N} - \mu)^2 = \mathbf{N}(\mathbf{N} - 1) + (1 - 2\mu)\mathbf{N} + \mu^2. \tag{3.10}$$

Now average over polynomials in $\Omega(q,d)$; take the expected value of both sides of (3.10), to get

$$\sigma^2 = \mathbb{E}_{\mathbb{M}}((\mathbf{N})_2) + (1 - 2\mu)\mu + \mu^2 = \mathbb{E}_{\mathbb{M}}((\mathbf{N})_2) + \mu - \mu^2. \tag{3.11}$$

To calculate μ, use Lemma 2.1 just like before in (3.5). (See also [9]). We chose β so that $d = 2\beta^2 \geq k$, therefore Lemma 2.1 applies. If $\mathcal{Z}_k$ denotes the set of all possible k-cycles, then $\mathbf{N} = \sum_{k \in J} \sum_{C \in \mathcal{Z}_k} I_C$, and

$$\mu = \sum_{k \in J} \sum_{C \in \mathcal{Z}_k} \mathbb{E}_{\mathbb{M}}(I_C) = \sum_{k \in J} |\mathcal{Z}_k| q^{-k} = \sum_{k \in J} \frac{(q)_k}{q^k} \frac{1}{k}.$$

Since $k \leq d = o(\sqrt{q})$, we get $\frac{(q)_k}{q^k} = \prod_{j=0}^{k-1}(1 - \frac{j}{q}) = 1 + o(1)$, just as in (3.7). Thus

$$\mu = (1 + o(1)) \sum_{k=\beta}^{\beta^2} \frac{1}{k} = (1 + o(1)) \log \beta. \tag{3.12}$$

As in Lemma 2.2, we have $(\mathbf{N})_2 = \sum_{C_1, C_2} I_{C_1} I_{C_2}$ where the sum is over distinct pairs of disjoint cycles C_1, C_2 whose lengths lie in J. If C_1 and C_2 are disjoint cycles of lengths k_1, k_2 respectively, then by Lemma 2.1, $\mathbb{E}_{\mathbb{M}}(I_{C_1} I_{C_2}) = q^{-k_1 - k_2}$. (We chose β so that $d = 2\beta^2 \geq k_1 + k_2$.) The number of ways to choose an ordered pair of disjoint cycles of length k_1 and k_2 is $\frac{q!}{k_1! k_2! (q - k_1 - k_2)!}(k_1 - 1)!(k_2 - 1)! = \frac{(q)_{k_1 + k_2}}{k_1 k_2}$ just as in the proof of Theorem 2.3. Hence

$$\mathbb{E}_{\mathbb{M}}((\mathbf{N})_2) = \sum_{k_1, k_2 \in J} \frac{(q)_{k_1 + k_2}}{q^{k_1 + k_2}} \frac{1}{k_1 k_2}.$$

Because $\beta^2 = o(\sqrt{q})$, we have $\frac{(q)_{k_1+k_2}}{q^{k_1+k_2}} = 1 + o(1)$ for all $k_1, k_2 \in J$. Therefore

$$\mathbb{E}_{\mathbb{M}}((\mathbf{N})_2) = (1 + o(1)) \sum_{k_1,k_2 \in J} \frac{1}{k_1 k_2} = (1 + o(1))(\log \beta)^2 = \mu^2(1 + o(1)).$$

Putting this back into (3.11), we get $\sigma^2 = o(\mu^2)$. By (3.9), this completes the proof. □

4 Rational Functions

If f and g are polynomials in $\mathbb{F}_q[x]$, let $\rho(g) =$ the degree of g and let $mgcd(f, g)$ be the greatest monic common divisor of f and g. Let

$$U(q, d) = \{(f, g) : \rho(f) = \rho(g) = d \text{ and } mgcd(f, g) = 1 \text{ and } g \text{ is monic}\}.$$

It is known [4], [7] that

$$|U(q, d)| = q^{2d+1}(1 - \frac{1}{q})^2. \tag{4.1}$$

For each pair $(f, g) \in U(q, d)$, the rational function $R(x) = \frac{f(x)}{g(x)}$ can be regarded as a function from $\mathbb{F}_q \cup \{\infty\}$ to $\mathbb{F}_q \cup \{\infty\}$ with the convention that $R(x) = \infty$ whenever $g(x) = 0$ and $R(\infty) =$ the leading coefficient of f. Note that, because of the conditions on f and g, the polynomials f and g are uniquely determined by the rational function R. Let $\mathbb{P}_d$ be the uniform probability measure on $U(q, d)$, and let $\mathbb{E}_d$ denote the corresponding expectation. Since $F_q \cup \{\infty\}$ is finite, the ultimate period is well-defined. Rather than introduce new notation, we reuse the same notation $\mathbf{T}$ for the ultimate period; in this section $\mathbf{T}$ defined on $U(q, d)$ instead of $\Omega(q, d)$.

Our goal in this section is to prove a lower bound for $\mathbb{E}_d(\log \mathbf{T})$. To avoid dealing with the exceptional point ∞, define $\hat{c}_k$ to be the number of cycles of length k that do not include ∞. Also let $\hat{\delta}_k = 1$ if $\hat{c}_k > 0$, $\hat{\delta}_k = 0$ otherwise. With this notation, $\mathbf{T}$ may be strictly larger than the least common multiple of the k's for which $\hat{c}_k > 0$. (It is possible that the only cycle of length k happens to contain ∞.) That poses no difficulties because we are proving lower bounds. As before, we have

$$\mathbf{T}(f) \geq \prod_{p \leq \zeta}^{*} p^{\hat{\delta}_p}, \tag{4.2}$$

where the $*$ indicates that the product is restricted to primes, and ζ is a parameter to be specified later.

Averaging over rational functions in $U(q, d)$, we get

$$\mathbb{E}_d(\mathbf{T}) \geq \mathbb{E}_d(\log \mathbf{T}) \geq \sum_{p \leq \zeta}^{*} \mathbb{E}_d(\hat{\delta}_p) \log p, \tag{4.3}$$

where the choice of the parameter ζ is postponed. Next we prove a rational function analogue of lemma 2.1.

Lemma 4.1. *Let $A \subset \mathbb{F}_q$, and let σ be a permutation of A. If $d \geq 2|A|$, then*

$$\mathbb{P}_d(\{(f,g) \in U(q,d) : \frac{f}{g} \text{ extends } \sigma\}) = q^{-|A|}\left(1 + O\left(\frac{|A|}{q}\right)\right).$$

Proof. From the definitions we have

$$\mathbb{P}_d(\{(f,g) \in U(q,d) : \frac{f}{g} \text{ extends } \sigma\}) =$$

$$\frac{1}{|U(q,d)|}\sum_g |\{f : \frac{f}{g} \in U(q,d) \text{ and } \frac{f(x)}{g(x)} = \sigma(x) \text{ for all } x \in A\}|, \tag{4.4}$$

where the sum is over monic degree d polynomials g that have no roots in A. As an auxiliary device, consider a superset of $U(q,d)$ in which the coprimality restriction is removed:

$$U^*(q,d) = \{(f,g) : \rho(f) = \rho(g) = d \text{ and } g \text{ is monic}\}.$$

If $(f,g) \in U^*(q,d)$, and $mgcd(f,g) = h$ has degree k, then $f^* := \frac{f}{h}$ and $g^* := \frac{g}{h}$ are coprime polynomials of degree $d-k$. Define

$$\Phi_1 = \frac{1}{|U(q,d)|}\sum_g |\{f : \frac{f}{g} \in U^*(d,q) \text{ and } \frac{f}{g} = \sigma(x) \text{ for all } x \in A\}|,$$

where, as in (4.4), the sum is over monic degree d polynomials g that have no roots in A. Comparing Φ_1 with (4.4), we see that Φ_1 is an overestimate because f is chosen from a larger set. We can therefore write

$$\mathbb{P}_d(\{R \in U(q,d) : R \text{ extends } \sigma\}) = \Phi_1 - \Phi_2 - \Phi_3, \tag{4.5}$$

where Φ_2 is the excess contribution from those pairs in $U^*(q,d) - U(q,d)$ for which the degree of the mgcd is is larger than one, but not larger than $|A|$, and Φ_3 is the remaining excess contribution from pairs with for which the *mgcd* has degree larger than $|A|$. To fit our equations on one line, abbreviate $U = |U(q,d)|$. Then we can be more explicit:

$$\Phi_2 = \frac{1}{U}\sum_{\{h:2\leq\rho(h)\leq|A|\}} |\{(f,g) : mgcd(f,g) = h, g \text{ monic, w.o. roots in } A, \frac{f}{g} \text{ extends } \sigma\}|,$$

and

$$\Phi_3 = \frac{1}{|U|}\sum_{\{h:\rho(h)>|A|\}} |\{(f,g) : mgcd(f,g) = h, g \text{ monic, w.o. roots in } A, \frac{f}{g} \text{ extends } \sigma\}|.$$

First we estimate the main term Φ_1. In Φ_1, the number of terms in the sum is the number of degree d monic polynomials with no root in A. We use inclusion-exclusion, with property i being that g has the i'th element of A as a root. For any particular set $\alpha_1, \alpha_2, \ldots \alpha_j$ of

roots, the number of monic degree d polynomials having those roots is the number that can written as $(x-\alpha_1)(x-\alpha_2)\cdots(x-\alpha_j)\tilde{g}(x)$ for some monic degree $d-j$ poly $\tilde{g}$, namely q^{d-j}. By inclusion-exclusion, the number of terms in the sum Φ_1, i.e. the number of degree d monic polynomials with no root in A, is

$$\sum_{j=0}^{|A|}\binom{|A|}{j}(-1)^j q^{d-j} = q^d(1-\frac{1}{q})^{|A|} = q^d(1+O(\frac{|A|}{q})).$$

Because $d > |A|$, we can use Lemma 2.1 again to calculate the terms in the sum Φ_1 for a given g:

$$|\{f : \rho(f) = d \text{ and } f(x) = \sigma(x)g(x) \text{ for all } x \in A\}| = q^{d-|A|}(q-1).$$

Combining our estimates for the size of terms, the number of terms, and the cardinality of $U(q,d)$, we get $\Phi_1 = \left(\frac{1}{q^{2d+1}(1-\frac{1}{q})^2}\right)\left(q^d(1-\frac{1}{q})^{|A|}\right)\left(q^{d-|A|}(q-1)\right)$. Thus

$$\Phi_1 = q^{-|A|}(1+O(\frac{|A|}{q})). \tag{4.6}$$

Next we show that Φ_2 is negligible compared to Φ_1 . If the degree of $h = mgcd(f,g)$ is k, then $f^* := f/h$ and $g^* := g/h$ have degree $d-k$. Given k, there are q^k choices of a monic polynomial $h = mgcd(f,g)$, and at most q^{d-k} choices of a monic polynomial g^* of degree $d-k$. We have $d \geq 2|A|$ as a hypothesis, and we have $k \leq |A|$ built into the definition of Φ_2. Therefore $d-k \geq |A|$ and Lemma 2.1 is applicable: there are, for a given g^*, at most $q^{d-k-|A|}$ choices for f^* such that $f^*(x) = \sigma(x)g^*(x)$ for all $x \in A$. Therefore

$$\begin{aligned}
\Phi_2 &\leq \frac{1}{|U(q,d)|}\sum_{k=2}^{|A|} q^k q^{d-k} q^{d-k-|A|} \\
&= \frac{q^{2d-|A|}}{q^{2d+1}(1-\frac{1}{q})^2}\sum_{k=2}^{|A|} q^{-k} \\
&= q^{-|A|}O(\frac{1}{q^3}) = O(\frac{\Phi_1}{q^3}).
\end{aligned}$$

Thus

$$\Phi_2 = O(\frac{\Phi_1}{q^3}). \tag{4.7}$$

We can be very crude about estimating Φ_3. We get a satisfactory bound by ignoring completely the requirement that $\frac{f^*}{g^*}$ extends σ, and g has no roots in A. We retain only the requirements that g and h are monic, and that $\rho(f^*) = \rho(g^*) = d - \rho(h)$. Thus

$$\Phi_3 \leq \frac{1}{|U(q,d)|}\sum_{\{\text{monic } h:\rho(h)>|A|\}} |\{(f^*,g^*) : \rho(f^*) = \rho(g^*) = d-\rho(h), g^* \text{ monic}\}|.$$

Given $k > |A|$, there are q^k choices of a monic polynomial h of degree k, q^{d-k} choices of a monic polynomial q^* of degree $d-k$, and $q^{d-k}(q-1) = q^{d-k+1}(1-\frac{1}{q})$ choices of a polynomial f^* of degree $d-k$. Therefore

$$\begin{aligned}\Phi_3 &\le \frac{1}{|U(q,d)|}\sum_{k=|A|+1}^{d} q^k q^{d-k} q^{d-k+1}(1-\frac{1}{q}) \\ &= \frac{1}{q^{2d+1}(1-\frac{1}{q})^2}\sum_{k=|A|+1}^{d} q^{2d+1-k}(1-\frac{1}{q}) \qquad = \frac{1}{(1-\frac{1}{q})}\sum_{k=|A|+1}^{d} q^{-k} = O\left(\frac{1}{q^{|A|+1}}\right).\end{aligned}$$

Thus

$$\Phi_3 = O\left(\frac{\Phi_1}{q}\right). \tag{4.8}$$

Putting (4.6), (4.7), and (4.8) into (4.5), we get lemma (4.1). □

With Lemma 4.1 in hand, the proof of a lower bound for $\mathbb{E}_d(\log \mathbf{T})$ is essentially the same as in the preceding section. We need a new notation for the rational function analogues of S_1 and S_2. Let

$$\hat{S}_r = \hat{S}_r(p) = \sum_{\{C_1,C_2,\ldots,C_r\}} \mathbb{P}_d\left(\bigcap_{i=1}^{r} A_{C_i}\right) = \sum_{\{C_1,C_2,\ldots,C_r\}} \mathbb{E}_d(\prod_{i=1}^{r} I_{C_i}),$$

(Cycles that contain ∞ are not included in this sum. As, before $\mathcal{Z}_p$ is the set of p-cycles that can be formed from the elements of $\mathbb{F}_q$.) Taking σ in Lemma 4.1 to be a p-cycle, we can estimate the sum $\hat{S}_1$. If $d \ge 2p$,

$$\begin{aligned}\hat{S}_1 &= \sum_{C\in\mathcal{Z}_p} \mathbb{P}_d(C \text{ is a cycle of } R) \\ &= \binom{q}{p}(p-1)!q^{-p}\left(1+O\left(\frac{p}{q}\right)\right).\end{aligned}$$

Thus

$$\hat{S}_1 = \frac{1}{p}\prod_{j=0}^{p-1}(1-\frac{j}{q})\left(1+O\left(\frac{p}{q}\right)\right) = \frac{1}{p}\prod_{j=0}^{p-1}(1-\frac{j}{q}) + O(\frac{1}{q}). \tag{4.9}$$

Similarly, if $d \ge 4p$, we can apply Lemma 4.1 to any permutation that consists of two p-cycles. Therefore

$$\begin{aligned}\hat{S}_2 &= \sum_{\{C_1,C_2\}} \mathbb{P}_d(C_1, C_2 \text{ are cycles}) \\ &= \frac{q(q-1)\cdots(q-2p+1)}{2p^2}q^{-2p}\left(1+O\left(\frac{p}{q}\right)\right)\end{aligned}$$

After simplification, this is

$$\hat{S}_2 = \frac{1}{2p^2}\prod_{j=0}^{2p-1}(1-\frac{j}{q}) + O(\frac{1}{pq}) \tag{4.10}$$

As in (3.4), we have

$$\hat{S}_1 - \hat{S}_2 \leq \mathbb{E}_d(\hat{\delta}_p) \leq \hat{S}_1. \tag{4.11}$$

Putting (4.9) and (4.10) into (4.11) and then (3.2), we get

$$\mathbb{E}_d(\log \mathbf{T}) \geq \sum_{p\leq\zeta}^{*}\frac{\log p}{p}\prod_{j=0}^{p-1}(1-\frac{j}{q}) - \sum_{p\leq\zeta}^{*}\frac{\log p}{2p^2}\prod_{j=0}^{2p-1}(1-\frac{j}{q}) + O(\sum_{p\leq\zeta}\frac{\log p}{q}).$$

If we take $\zeta = \frac{d}{4}$, and require $d = o(\sqrt{q})$, then $\sum_{p\leq\zeta}\frac{\log p}{q} = o(1)$ and, as in (3.7), the products are $1 + o(1)$. We therefore have a rational function analogue of Corollary 3.2:

Theorem 4.2. *If $d = d(q) \to \infty$ is such a way that $d = o(\sqrt{q})$, then*

$$\mathbb{E}_d(\log \mathbf{T}) \geq \frac{d}{4}(1 + o(1)).$$

Acknowledgement: We thank Igor Shparlinski and BIRS, for providing reference [9] and stimulating our interest in this subject. We also thank Derek Garton for helpful comments. A referee's suggestions helped us improve both the proofs and the exposition.

References

[1] Andrews, George E., Number theory, (Corrected reprint of the 1971 original) Dover Publications, Inc., New York, 1994, ISBN 0-486-68252-8

[2] Bach, Eric, Toward a theory of Pollard's rho method, Inform. and Comput., *Information and Computation*, **90**, (1991), no.2, 139–155.

[3] Benedetto, Robert L. and Ghioca, Dragos and Hutz, Benjamin and Kurlberg, Pär and Scanlon, Thomas and Tucker, Thomas J., Periods of rational maps modulo primes, *Math. Ann.*, **355**, (2013), no.2,637–660.

[4] Benjamin, Arthur T. and Bennett, Curtis D., The probability of relatively prime polynomials, *Math. Mag.*, **80**, 2007, no.3, 196–202.

[5] Billingsley, Patrick, Probability and measure,second edition, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986, ISBN 0-471-80478-9.

[6] Bollobás, Béla, Random graphs, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1985, ISBN = 0-12-111755-3; 0-12-111756-1.

[7] Corteel, Sylvie and Savage, Carla D. and Wilf, Herbert S. and Zeilberger, Doron, A pentagonal number sieve, *J. Combin. Theory Ser. A*, **82** (1998), no. 2, 186–192.

[8] Flajolet, Philippe and Odlyzko, Andrew M., Random mapping statistics, in Advances in cryptology—EUROCRYPT '89 (Houthalen, 1989), Lecture Notes in Comput. Sci., **434**, 329–354, Springer, Berlin, 1990.

[9] Flynn, Ryan and Garton, Derek, Graph components and dynamics over finite fields, *Int. J. Number Theory*, **10**, 2014, no. 3, 779–792.

[10] Harris, Bernard, Probability distributions related to random mappings, *Ann. Math. Statist.*, 1960, 1045–1062.

[11] Harris, Bernard, The asymptotic distribution of the order of elements in symmetric semigroups, *J. Combinatorial Theory Ser. A*, **15**, (1973), 66–74.

[12] Hua, Loo Keng, Introduction to number theory, Translated from the Chinese by Peter Shiu, Springer-Verlag, Berlin-New York, 1982.

[13] Janson, Svante and Łuczak, Tomasz and Rucinski, Andrzej, Random graphs, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000, ISBN 0-471-17541-2.

[14] Kolchin, Valentin F., Random mappings, Translation Series in Mathematics and Engineering Optimization Software, Inc., Publications Division, New York, 1986,ISBN 0-911575-16-2,

[15] Landau, Edmund, Handbuch der Lehre von der Verteilung der Primzahlen. 2 Bände,2d ed, Chelsea Publishing Co., New York, 1953.

[16] Martins, Rodrigo S. V. and Panario, Daniel, On the Heuristic of Approximating Polynomials over Finite Fields by Random Mappings, arXiv:1505.02983.

[17] Massias, J.-P. and Nicolas, J.-L. and Robin, G., Évaluation asymptotique de l'ordre maximum d'un élément du groupe symétrique, *Acta Arith.*, **50**, 1988, no. 3, 221–242.

[18] Pollard, J. M., A Monte Carlo method for factorization, *Nordisk Tidskr. Informationsbehandling (BIT)*, **15**, (1975), no.3, 331–334.

[19] Schmutz, Eric, Period lengths for iterated functions, *Combin. Probab. Comput.*, **20**, 2011, no. 2, 289–298.

[20] Silverman, Joseph H., Variation of periods modulo p in arithmetic dynamics, *New York J. Math.*, **14**, (2008), 601–616.

[21] Silverman, Joseph H., The arithmetic of dynamical systems, Graduate Texts in Mathematics, **241**, Springer, New York, (2007), ISBN 978-0-387-69903-5.

[22] Stanley, Richard P., Enumerative combinatorics. Vol. I, The Wadsworth & Brooks/Cole Mathematics Series, 1986 0-534-06546-5.